A

NEW, CONTINUATION, DIVISIONAL OR CONTINUATION-IN-PART APPLICATION UNDER 37 C.F.R. §1.53(b)

Attorney Docket No. 9432-000089

Express Mail Label No. EJ 179 205 142 US

Date October 28, 1999

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Hon. Commissioner of Patents and Trademarks
Washington, D. C. 20231

Sir:

Transmitted herewith for filing under 37 C.F.R §1.53(b) is a patent application for

Content-Based Graph Authentication of Graph Presented in Text Documents

identified by: [ ] First named inventor ______
or [X] Attorney Docket No. (see above)

1. **Type of Application**

[X] This application is a new (non-continuing) application.

[ ] This application is a [ ] continuation / [ ] divisional / [ ] continuation-in-part of prior application No. _____. Amend the specification by inserting before the first line the sentence:

> --This is a [continuation/division/continuation-in-part] of United States patent application No. _____, filed _____.--

[ ] The entire disclosure of the prior application, from which a copy of the oath or declaration is supplied, is considered part of the disclosure of the accompanying application and is hereby incorporated by reference therein.

If for some reason applicant has not requested a sufficient extension of time in the parent application, and/or has not paid a sufficient fee for any necessary response in the parent application and/or for the extension of time necessary to prevent the abandonment of the parent application prior to the filing of this application, please consider this as a Request for an Extension for the required time period and/or authorization to charge our Deposit Account No. 08-0750 for any fee that may be due. THIS FORM IS BEING FILED IN TRIPLICATE: one copy for this application; one copy for use in connection with the Deposit Account (if applicable); and one copy for the above-mentioned parent application (if any extension of time is necessary).

2. **Contents of Application**

a. Specification of **28** pages;
[ ] A microfiche computer program (Appendix);
[ ] A nucleotide and/or amino acid sequence submission;

[ ] Because the enclosed application is in a non-English language, a verified English translation [ ] is enclosed [ ] will be filed.

[ ] Cancel original claims _____ of the prior application before calculating the filing fee. (At least one original independent claim must be retained for filing date purposes.)

b. [X] Drawings on **12** sheets;

c. [X] A signed Declaration [X] is enclosed / [ ] will be filed in accordance with 37 C.F.R. §1.53(f).

The enclosed Declaration is [X] newly executed / [ ] a copy from a prior application under 37 C.F.R. §1.63(d) / [ ] accompanied by a statement requesting the deletion of person(s) not inventors in the continuing application.

d. **Fees**

FILING FEE	Number				Number				Basic Fee
CALCULATION	Filed				Extra		Rate		$760.00
Total Claims	33	-	20	=	13	×	$18.00	=	234.00
Independent Claims	3	-	3	=	0	×	$78.00	=	
Multiple Dependent Claim(s) Used							$260.00	=	
FILING FEE - NON-SMALL ENTITY									994.00
FILING FEE - SMALL ENTITY: Reduction by 1/2 .... [ ] Verified Statement under 37 C.F.R. §1.27 is enclosed. [ ] Verified Statement filed in prior application.									
Assignment Recordal Fee ($40.00)									
37 C.F.R. §1.17(k) Fee (non-English application)...									
TOTAL									**994.00**

[X] A check is enclosed to cover the calculated fees. The Commissioner is hereby authorized to charge any additional fees that may be required, or credit any overpayment, to Deposit Account No. 08-0750. A duplicate copy of this document is enclosed.

[ ] The calculated fees will be paid within the time allotted for completion of the filing requirements.

[ ] The calculated fees are to be charged to Deposit Account No. 08-0750. The Commissioner is hereby authorized to charge any additional fees that may be required, or credit any overpayment, to said Deposit Account. A duplicate copy of this document is enclosed.

3. **Priority Information**

[ ] **Foreign Priority**: Priority based on ______ Application No. _______, filed ______, is claimed.

[ ] A copy of the above referenced priority document [ ] is enclosed / [ ] will be filed in due course, pursuant to 35 U.S.C. §119(a)-(d).

[ ] **Provisional Application Priority**: Priority based on United States Provisional Application No. _____, filed _____, is claimed under 35 U.S.C. §119(e).

Attorney Docket No. 9432-000089

Express Mail Label No. EJ 179 205 142 US

Date October 28, 1999

4. **Other Submissions**

[ ] A Preliminary Amendment is enclosed.

[ ] An Information Disclosure Statement, ____ sheets of PTO Form 1449, and ____ patent(s)/publications/documents are enclosed.

[X] A power of attorney

[X] is submitted [X] with the new Declaration.

[ ] is of record in the prior application and [ ] is in the original papers / [ ] a copy is enclosed.

[ ] An Assignment of the invention

[ ] is enclosed with a cover sheet pursuant to 37 C.F.R. §§3.11, 3.28 and 3.31.

[ ] is of record in a prior application. The assignment is to ____, and is recorded at Reel ____, Frame(s) ____.

[ ] An Establishment of Assignee's Right To Prosecute Application Under 37 C.F.R. §3.73(b), and Power Of Attorney is enclosed.

[X] An Express Mailing Certificate is enclosed.

[X] Other: return postcard

Attention is directed to the fact that the correspondence address for this application is:

Harness, Dickey & Pierce, P.L.C.
P.O. Box 828
Bloomfield Hills, Michigan 48303
(248) 641-1600.

Respectfully,

Greg Stobbs

Gregory A. Stobbs
Reg. No. 28764

Date: October 28, 1999
Harness, Dickey & Pierce, P.L.C.
P.O. Box 828
Bloomfield Hills, Michigan 48303
(248) 641-1600

HARNESS, DICKEY & PIERCE, P.L.C.
ATTORNEYS AND COUNSELORS
P O. BOX 828
BLOOMFIELD HILLS, MICHIGAN 48303
U.S.A.

TELEPHONE
(248) 641-1600

TELEFACSIMILE
(248) 641-0270

Date October 28, 1999

Hon. Commissioner of Patents
and Trademarks
Washington, D.C. 20231

Sir:

EXPRESS MAILING CERTIFICATE

Applicant: Hong Heather Yu

Serial No (if any):

For: **Content-Based Authentication of Graph Presented in Text Documents**

Docket: 9432-000089

Attorney: Gregory A. Stobbs

"Express Mail" Mailing Label Number............................EJ 179 205 142 US

Date of Deposit ....October 28, 1999

I hereby certify and verify that the accompanying **return postcard; $994.00 check ($760 for filing fee and $234 for extra claims); 3-page transmittal letter (in triplicate); 28-page application; 12 sheets of drawings (showing Figures 1-18); 2-page Declaration and Power of Attorney; along with copies of this Express Mailing Certificate** are being deposited with the United States Postal Service "Express Mail Post Office To Addressee" service under 37 C.F.R. 1.10 on the date indicated above and are addressed to the Commissioner of Patents and Trademarks, Washington, D.C. 20231.

Pamela Strauss
Signature of Person Mailing Documents

CONTENT-BASED AUTHENTICATION OF GRAPH PRESENTED IN TEXT DOCUMENTS

BACKGROUND OF THE INVENTION

Technical Field

The present invention relates generally to document authentication. More particularly, the present invention relates to the authentication of graphs at the object level as well as the pixel level.

Discussion

For as long as humans have communicated with one another, there has been concern over maintaining confidentiality. As a result, verbal, written, and electronic messages have all been the subject of substantial technological efforts to maintain security. For example, document authentication techniques are commonly used to ensure the integrity of a wide variety of electronic documents such as, presentations, contracts, military orders, and databases. Authentication involves the task of making the determination that the document has not been tampered with and that it originated with the presumed transmitter. Authentication using digital watermarks is a particular technique that has been studied by many researchers in the last ten years. For example, digital watermarking has been successfully applied to digital documents such as digital color/gray scale images and plain text. While electronic document authentication efforts have experienced considerable success, it is important to note that these efforts have typically centered around the protection of textual documents and images.

Recently, however, more and more documents are using graphs in addition to images and text for system and idea illustration. In contrast to images, graphs are more difficult to watermark because of low capacity of additive noise. This is due to the binary nature of graphs. The term "binary nature" relates to the fact that most graphs have one bit per pixel, whereas most images have multiple bits per pixel to indicate varying shades and colors. Binary pixels make it particularly hard to insert watermarks due to the low capacity for perceptual invisible noise. In other words, a minimal alteration of bits in a binary graph can result in a substantial change in the appearance and content of the graph. Furthermore, the critical information of a graph is often contained at the object level rather than the pixel level. For example, a useful application for document copying and copyright protection is to provide different levels of access to different users. In such a case it would be very desirable to detect alteration of the original document as well as localize the alteration on the object level. For example, it is more important to detect a substantive change in a document, such as "10%" to "70%", than it is to detect an increase in the size of an arrow by one pixel. Thus, the sensitive information in a document is often contained on the object level rather than the pixel level.

Pixel level authentication may also result in less flexibility. For example, if the annotation font of a graph changes but the content of the graph does not, pixel level authentication will alert the owner that the annotations have been altered. The owner has no way of determining, however, that the content of the graph matches the original. Object level authentication, on the other hand, would assure the owner that the "content is authentic" in such a case. If the font is marked as sensitive information, object level authentication could also alert the owner to font alterations. In many applications, however, it would be highly desirable to provide a mechanism for returning an "authentic" determination if the font is not marked as sensitive information.

Conventional methodologies for content-based text authentication mainly rely on altering the word/line spacing or the length of character vertical serif strokes. While text documents are often referred to as binary images and share the same binary nature of graphs, these methodologies can hardly be extended to authentication of graphs. This is because even on the pixel level graphs generally do not exhibit the same characteristics as text. For instance, in a graphical flowchart the shape of each node may be very important, whereas the nodes often have substantially fewer characters as compared to a paragraph of text. In such a flowchart the number of objects that exhibit a vertical serif can be as low as a few percent of the total number of objects. Here, an object is referred to an alterable line, character, or curve.

In fact, other kinds of graphs may not exhibit alterable line spacing or vertical serif at all. It is therefore desirable to bridge text-based authentication techniques to the authentication of graphs.

SUMMARY OF THE INVENTION

The above and other objects are provided by a computerized method for authenticating a document. The method includes the step of partitioning the document into graphical content and textual content. The graphical content is then converted into a symbolic representation of the graphical content. The method further provides for authenticating the symbolic representation with a text authentication algorithm.

The present invention also provides a computerized method for authenticating a binary graph. The graph is authenticated at the pixel level as well as the object level. The method includes the step of encrypting the authenticated graph.

As a further aspect of the invention, a graph authentication system has an object level authenticator for authenticating a graph at an object level. The authentication system further includes a pixel level authenticator for authenticating the graph at a pixel level and an encryption system for encrypting the authenticated graph.

It is to be understood that both the foregoing general description and the following detailed description are merely exemplary of the invention, and are intended to provide an overview or framework for understanding the nature and character of the invention as it is claimed. The accompanying drawings are included to provide a further understanding of the invention, and are incorporated in and constitute part of this specification. The drawings illustrate various features and embodiments of the invention, and together with the description serve to explain the principles and operation of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

The various advantages of the present invention will become apparent to one skilled in the art by reading the following specification and appended claims, and by referencing the following drawings in which:

Figure 1 is a block diagram of a graph authentication system according to the present invention;

Figure 2 is a block diagram of an object level authenticator according to the present invention;

Figure 3 is a block diagram of a pixel level authenticator according to the present invention;

Figure 4 is a flowchart of a computerized method for authenticating a document according to the present invention;

Figure 5 is a flowchart of the process of authenticating a graph at the object level according to the present invention;

Figure 6 is a flowchart of the process for authenticating a graph at the pixel level according to the present invention;

Figure 7 is a flowchart for the process of adding visible authorization information according to the present invention;

Figure 8 is a flowchart for the process of adding invisible authorization information according to the present invention;

Figure 9 is a sample illustration of graphs which can be authenticated with the present invention;

Figure 10 is an illustration of a graphical flowchart which can be authenticated with the present invention;

Figure 11 is a block diagram of a one-party owned document authentication process according to the present invention;

Figure 12 is an illustration of a key set according to the present invention;

Figure 13 is a table of relationship and specification symbols according to a preferred embodiment of the present invention;

Figure 14 is a symbolic representation of the graphical flowchart of Figure 10;

Figure 15 is a graphical flowchart authenticated at the pixel level using a bounded box according to the present invention;

Figure 16 is a graphical flowchart authenticated at the pixel level using a bar code according to the present invention;

Figure 17 is an enlarged view of textual and graphical content containing invisible authentication information; and

Figure 18 is a table comparing graph authentication algorithms.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

Turning now to Figure 1, the preferred embodiment of the graph authentication system 20 includes an object level authenticator 30, a pixel level authenticator 40, and an encryption system 50. The graph authentication system 20 provides for content-based authentication of graphs contained in a host document 51. The result is protected document 52. As part of the following discussion, I is defined to be the host document 51, such as a contract, which will be authenticated by owner O1 or owners O1, O2 to On. The authenticated copy of host document I is denoted as $\tilde{I}$. In correspondence, G and $\tilde{G}$ are defined to be the original and the authenticated copy of a graph respectively. Furthermore, R is defined as an authorized receiver, whereas A is an attacker, i.e., unauthorized receiver. The following scenarios illustrate potential applications and objectives of graph authentication system 20.

The first scenario is the situation in which $I_1 \in$ O1, O1 wants to determine whether her document I_1 is authentic. The content of the document contains sensitive information, such as a price of $1,000 or a deadline of June 01, 1999. Another scenario occurs when $I_1 \in$ O1, O1 needs to send I_1 to R and wishes to grant R "read" permission but not "write" permission. A variation on this scenario is the situation in which O1 wants to prevent alteration of any kind and to localize the alterations made by an attacker A who gets I_1 from O1 and then sends it to R. Or, O1 may want everyone to be able to read I_1 while only herself and R can make modification on the document. Another scenario occurs when $I_1 \in O1 \cap O2$, i.e., I_1 is a contract between O1 and O2. If the copy in O1's hand is different from that of in O2's, O1 wants to prove that O2's copy is a tampered copy of the original contract by checking the authenticity of O2's copy. In addition, O1 may want to point out where exactly O2 altered the original contract.

Turning now to Figure 11, it can be appreciated that the present invention provides a fully functional content-based authentication system for text documents including binary graphs. By building a bridge from graph to text on the character level, the present invention allows authentication of graphs using suitable text document authentication algorithms. When pixel level precision of a graph is required, a pixel level authentication can be added. This layer lets the owner detect as well as localize changes in the graph on the pixel level. The hierarchical layout allows the application of the present invention to the aforementioned scenarios as well as other scenarios.

The first level of the hierarchy is the pixel level authentication which is followed by an object level authentication. These are done with owner *O1*'s private key. Notice here, either the pixel level or the object level protection is optional depending on the application. For ultimate protection, however, a dual-layer protection with a pixel protection layer plus an object protection layer is recommended since the two layers are orthogonal. Additionally, a meaningful watermark, such as a company logo, can be inserted, if desirable, into the document. Furthermore, the authenticated documents, including text and graphs can be encrypted with a public key encryption algorithm for secure transmission. Here the watermarking layer can be done either before or after the authentication layer. This again, depends on different applications. Access authorization can then be granted by distributing different keys to different users. For example, in the case of "read" only access, *R* will be given the public decryption key K_4 only. In the case of a multi-party owned document authentication, each party has a private key, the authentication is done by generating a key set with the private key from every party (see Figure 12). Attempted modifications of the document without a key will therefore be unsuccessful.

Returning to Figure 1, it will be appreciated that the object level authenticator 30 authenticates the graph at an object level, whereas the pixel level authenticator 40 authenticates the graph at a pixel level. The encryption system 50 encrypts the authenticated graph for transmission to the recipient. As seen in Figure 2, it will be appreciated that the object level authenticator 30 converts the graph into a symbolic representation of the graph via a specification module 31 and a relationship module 32. The specification module 31 defines nodes of the graph with specification symbols. Similarly, the relationship module

32 defines relationships between the nodes of the graph with relationship symbols. This allows a text authentication module 33 to authenticate the symbolic representation with a text authentication algorithm.

Figure 9 demonstrates the various types of graphs which can be authenticated via the present invention. The operation of the object level authenticator 30 can be better understood through the graphical flowchart of Figure 10. It can be appreciated that the important information contained in graphical flowchart 34 is the annotation of each node and the connections between nodes that illustrate the relationship of nodes. Whether the drawing of each box is slightly smaller or slightly larger, the length of a line is longer or shorter, or the position of a node is tilted to the left or right is generally not as important. Consequently, the authentication process is mainly concerned with the object level instead of the pixel level of the graphical flowchart 34. It is important to note that the important characteristics of an object depend on the type of graph. Thus, in the case of the bar chart of Figure 9(c), the important information is contained in the relative height of each individual bar rather than the overall height of the graph. For example, if the height of the second bar is changed to half its original height, the value of the second bar is thereby altered. It will be appreciated that the concern with most text documents is at the object level, or character level.

Graphical flowchart 34 therefore includes various nodes and lines and can be represented with a series of relationship symbols along with the node annotations as follows: "<N_1{'Process A', #1, &reg, @mid} → N_2{'Process B', #1, &reg, @mid}→ N_3{'If C', #3, &reg, @mid}→< N_4{'End', #2, &reg, @mid}|yes; N_2|no>>", wherein Figure 13 illustrates the relationship and specification symbols for the above symbolic representation. The result is shown in Figure 14. In the above symbolic representation, N_1 N_2... are node names with the property of each node contained in {}, < > is a tuple, and → and | are relationship symbols. It will be appreciated that the properties of nodes and lines, the shape, size, color, and position, can be described with the specification symbols. For those specification insensitive graphs, the symbols between each pair of {} can be simply ignored whereas in specification sensitive graphs, the specification symbols in each pair of {} provide different levels of details. This hierarchical representation provides additional flexibility.

After defining the nodes of the graph with specification symbols, and the conditions and familial relationships with relationship symbols, the text authentication module 33 can authenticate the symbolic representation. For example, well known two- or multi-dimensional checksum techniques can be used to verify authenticity. For the following discussion, let T(p,q) represent the (p,q)th character. $S(p,q) = s^1(p,q)\ s^2(p,q)\ ^{\cdots}s^J(p,q) = f(T(p,q))$ is the coded representation of T(p,q) via map f, wherein $s^1(p,q)\ s^2(p,q)\ ^{\cdots}s^J(p,q)$ represent the first, the second, ... and the Jth bit of S(p,q) that are in the order of the most significant bit to the least significant bit. Furthermore, let $Sum_p^j= \sum^P{}_{p=1}$

$s^j(p,q)$ and $Sum_q^j = \sum_{q=1}^{Q} s^j(p,q)$, where P & Q are dimensional sizes. Thus, the position (p,q) of any alteration $Sum_p' \neq Sum_p$, $Sum_q' \neq Sum_q$ can be localized.

It will be appreciated that utilizing well known content-dependent one way hash functions provides a higher level of security. For the following discussion, let B denote the block size and K denote a private key. In the case of a multi-party document, K is a function of K_{O1}, K_{O2}, ..., i.e., K = f1(K_{O1}, K_{O2}, ...). Figure 12 illustrates a key set for the present example. For the purpose of discussion, we may assume each key in the set, K_{o1}, K_{o2}, ... to be encrypted with its owner's private key, and an arbitrator (a trusted third party) is used to generate the key set K. It is important to note, however, that other suitable cryptography protocols may also be used. Assume K is a Jbits coding with the 1st to (J-1)th bits being the code bits and the lowest bit, Jth bit, being the verification bit. The document paragraph I shown in Figure 14 can use 9bits coding. Choosing the one way hash algorithm MD5, the encoding procedure is as follows. Pad the source text I to an exact multiple of 512 in length. For each 128-length set, I_o, choose its neighborhood set, $\underline{I}_o$=512 characters with $I_o \subset \underline{I}_o$. Assume

$S_o = \{S_o(i), i \in [1,128]\} = \{s^1{}_o(i)\ s^2{}_o(i) \cdots s^J{}_o(i)\} = f(I_o)$

and

$\underline{S}_o = \{\underline{S}_o(i), i \in [1,512]\} = \{\underline{s}^1{}_o(i)\ \underline{s}^2{}_o(i) \cdots \underline{s}^J{}_o(i)\} = f(\underline{I}_o)$

are coded representation of I_o and $\underline{I}_o$ respectively.

1. Concatenate the code bits of the neighborhood set $\underline{I}_o$,

2. Calculate the 128bits hash value of it, $h_o=H(\underline{S_o})$,

3. Generate message $h_o'=Sgn(K,\ h_o)$ by signing h_o with public cryptography method, and

4. Put h_o' into the J^{th} bit, the lowest bit, of $S_o(i)$, i.e., let $s^J_o(i)= h_o'(i)$, $i\in[1,128]$.

The above algorithm is discussed in the context of image authentication in the article "Fragile imperceptible digital watermark with privacy control", C. W. Wu, D. Coppersmith, F. C. Mintzer, C. P. Tresser, and M. M. Yeung, IS&T/SPIE Conference on Security and Watermarking of Multimedia Content, SPIE 3657, Jan, 1999, incorporated herein by reference. The decoding process is similar to the encoding process with the verification done through an XOR operation. Such that $Auth_o(i)=\tilde{h}_o'(i)\oplus s^J_o(i)$.

If $Auth_o(i)=1$ for $\forall i\in[1,128]$, the I_o set has been altered.

Turning now to Figure 3, the pixel level authenticator 40 of the graph authentication system 20 is shown in greater detail. It can be appreciated that a visible watermarking module 41 adds visible authentication information to the graph at the pixel level, whereas an invisible watermarking module 42 adds invisible authentication information to the graph at the pixel level. The preferred embodiment further includes a coalescing module 43 for embedding a hash value from the object level of the graph at the pixel level of the graph. Dual level authentication with coalescing has been found to yield optimum results. To authenticate I with N symbols, we compute the one way hash of I on the character level first. Therefore, if N=248 characters this is done by putting all the bits of the 248 characters together, pad the result to an exact multiple of 512 in length, and calculate the hash value of the padded message. Then, the 128bits hash value is embedded at the pixel level.

Operation of the graph authentication system of the present invention will now be described in greater detail for programming purposes. Turning to Figure 4, a computerized method for authenticating an electronic file (or document) is shown generally at 100. Step 102 demonstrates receipt of the electronic file. At step 101, the file is partitioned into graphical content and textual content. The partitioning of graphs from text regions in a document has been the subject of considerable study. For example, U.S. Patent No. 5,465,304, and U.S. Patent No. 5,335,290 to Cullen, et al., incorporated herein by reference, discuss the segmentation of text, pictures, and lines of a document image. Furthermore, U.S. Patent No. 5,073,953 to Westdijk, incorporated herein by reference, discloses a system and method for automatic document segmentation. The separation of body text from other regions of a document is taught in U.S. Patent No. 5,892,843 to Zhou, et al., incorporated herein by reference. Also, in U.S. Patent No. 5,379,130 to Wang, et al., a method and system that separates images from text is disclosed. Any of these techniques or other well known approaches can be readily adapted to perform partitioning step 101.

At step 110, it is determined whether the object level is a level of concern. If so, the graph is authenticated at the object level at step 111 by adding authentication information the electronic file based on an object level representation. Similarly, at step 130 it is determined whether the pixel level is a level of concern. If so, the document is authenticated at the pixel level at step 131. It will be appreciated that object level authentication and pixel level authentication are both optional and can be performed in any order. The graph can then be encrypted at step 150 and transmitted at step 160 to an authorized recipient.

Figure 5 shows step 111 in greater detail. It can be appreciated that nodes of the graph are defined with specification symbols at step 112. Relationships between the nodes are then defined with relationship symbols at step 113. The result is a symbolic (or object level) representation of the graphical content contained in the electronic file. It will be appreciated that other approaches to object level representation can be taken without parting form the scope of the invention. At step 114, the symbolic representation is authenticated with a text authentication algorithm.

Turning now to Figure 6, step 131 is shown in greater detail. At step 132, it is determined whether transparency is required based on the content of the graph and the host document. If so, invisible authorization information is added at step 133. Otherwise, visible authorization information can be added at step 134.

As seen in Figure 7, a relatively robust approach for adding visible authorization information is shown in greater detail. Specifically, at step 135 a truncated image of the graph is formed. For the following discussion, let XxY=128 be the defined block size. Graph G can therefore be cut into XxY blocks. Assuming the number of blocks is L, we concatenate the bits of the (x,y)th pixel of every block to the 1st block and form an Lbits truncated image TrunG. Therefore, a Lbits/pixel image TrunG, with image size XxY, of graph G is generated. Let $\text{TrunG}(x,y)^{\ell}$ denote the ℓ^{th} bit of pixel (x,y) of TrunG. Notice here, it is desirable to form the truncated image TrunG in such a way that $\text{TrunG}(x,y)\neq 0$. Also note that to get a higher level of protection, a random number generator should be used to cut the graph.

At step 136, an initial message is generated from the truncated image. The initial message is defined by all bits of the truncated image. Thus, step 136 collects all bits of all XxY pixels into a XxYxL bits message M1. At step 137, the initial message is converted into a padded message, wherein the padded message has a size defined by a multiple of a predetermined length. Thus, M1 is padded into an exact multiple of 512 in length with as many zeros as needed to obtain message M1'.

At step 138, a hash value for the padded message is computed. Thus, step 138 computes the 128 bits hash value of M1' using MD5, M2=h(I)=H(M1'). At step 139, the hash value is converted into a public key encrypted message by signing the hash value with a public key cryptography method such that M3=h'(i)=Sgn(K, M2). The public key encrypted message is then converted into visible authentication information at step 140. The visible authentication information can be in many different formats. For example, Figure 15 illustrates an authenticated graph using a bounding box, whereas Figure 16 illustrates an authenticated graph using a bar code.

When invisible authentication is required or desirable, a less robust scheme that modifies the graph itself can be used. Thus, as shown in Figure 8, a truncated image is formed from the graph at step 135'. At step 141, a verification bit is selected from each pixel of the truncated image. Thus, at step 141 1bit $TrunG(x,y)^l$ =1 out of the Lbits of each pixel (x,y) in TrunG to be the verification bit. For better imperceptibility and a higher lever of security, the verification bits should be picked in a way to maximize spread.

At step 136' an initial message is generated from the truncated image, wherein the initial message is defined by all non-verification bits of the truncated image. Step 136' therefore collects the remaining (L-1) bits of all XxY pixels into a XxYx(L-1) bits message M1. Message M1 is padded into an exact multiple of 512 in length with as many 0s as needed and get message M1'. The initial message is therefore converted into a padded message at step 137'. Preferably, the padded message has a size defined by a multiple of a predetermined length of 512.

At step 138', the hash value is computed for the padded message. The hash value is then converted into a public key encrypted message at step 139'. The public key encrypted message can then be imbedded into the truncated image at step 142 in the following fashion:

- If h'(i)=h'((y-1)*X+x))=0 and |TrunG(x,y)|= odd, let TrunG(x,y)'=0.
- If h'(i)=h'((y-1)*X+x))=1 and |TrunG(x,y)|= even, let TrunG(x,y)'=1.

Where |TrunG(x,y)| denotes the cardinality of TrunG(x,y), i.e., the number of bits that are '1's among the Lbit of TrunG(x,y).

Turning now to Figure 17, two sample results can be seen. The lower result is cropped from the graphical flowchart in Figure 10. To give a better view, each result is enlarged to at least 400 percent of the original size.

Conventional space-shifting methods and serif-modification methods are proposed in "Electronic Marking and Identification Techniques to Discourage Document Copying", J. Brassil, S. Low, N. Maxemchuk, and L. O'Gorman, IEEE Infocom 94, and in "Document Marking and Identification using Both Line and Word Shifting", S. H. Low, N. F. Maxemchuk, J. T. Brassil, and L. O'Gorman, Infocom '95, both incorporated herein by reference. Comparing these techniques to the present invention, it can be seen in Figure 18 that clear improvement has been achieved. Notice that when the hash value is prepended to the document, special coding is not needed for object level authentication. Otherwise, such coding is needed. Similarly, in the case of pixel level or coalesced authentication, special coding is not needed with visible authentication information, whereas it is needed for invisible authentication information. Here, special coding means a new code other than commonly accepted codes, such as ASCII Code and Unicode.

The foregoing discussion discloses and describes exemplary embodiments of the present invention. One skilled in the art will readily recognize from such discussion, and from the accompanying drawings and claims, that various changes, modifications and variations can be made therein without departing from the spirit and scope of the invention as defined in the following claims.

WHAT IS CLAIMED:

1. A computerized method for authenticating an electronic file, the method comprising the steps of:

receiving an electronic file having a graphical content;

generating an object level representation of the graphical content; and

adding authentication information to the electronic file based on the object level representation of the graphical content.

2. The method of claim 1 wherein the graphical content contains binary pixel bit values.

3. The method of claim 1 further comprising the step of converting the graphical content into a symbolic representation of the graphical content.

4. The method of claim 3 further comprising the steps of:

defining nodes of the graphical content with specification symbols; and

defining relationships between the nodes of the graphical content with relationship symbols.

5. The method of claim 4 further comprising the step of defining the shape, size, color, and position of the nodes.

6. The method of claim 4 further comprising the step of defining conditions and familial relationships between the nodes.

7. The method of claim 1 further comprising the step of authenticating the object level representation with a text authentication algorithm.

8. The method of claim 7 further comprising the step of authenticating the object level representation with a checksum.

9. The method of claim 8 wherein the checksum is a two-dimensional checksum.

10. The method of claim 8 wherein the checksum is a multi-dimensional checksum.

11. The method of claim 7 further comprising the step of authenticating the object level representation with a cryptographic hash function.

12. The method of claim 1 further comprising the step of authenticating the graphical content at a pixel level.

13. The method of claim 12 further comprising the step of adding visible authentication information to the graphical content.

14. The method of claim 13 wherein the visible authentication information includes a bounding box.

15. The method of claim 13 wherein the visible authentication information includes a bar code.

16. The method of claim 12 further comprising the step of adding invisible authentication information to the graphical content.

17. The method of claim 1 further comprising the step of partitioning the electronic file into graphical content and textural content.

18. A computerized method for authenticating a binary graph, the method comprising the steps of:

authenticating the graph at a pixel level;

authenticating the graph at an object level; and

transmitting the authenticated graph to a recipient.

19. The method of claim 18 further comprising the step of adding visible authentication information to the graph.

20. The method of claim 19 further comprising the steps of:

forming a truncated image from the graph;

generating an initial message from the truncated image, the initial message defined by all bits of the truncated image;

converting the initial message into a padded message, the padded message having a size defined by a multiple of a predetermined length;

computing a hash value for the padded message;

converting the hash value into a public key encrypted message; and

converting the public key encrypted message into the visible authentication information.

21. The method of claim 20 wherein the visible authentication information includes a bounding box.

22. The method of claim 20 wherein the visible authentication information includes a bar code.

23. The method of claim 18 further comprising the step of adding invisible authentication information to the graph.

24. The method of claim 23 further comprising the steps of:

forming a truncated image from the graph;

selecting a verification bit from each pixel of the truncated image;

generating an initial message from the truncated image, the initial message defined by all non-verification bits of the truncated image;

converting the initial message into a padded message, the padded message having a size defined by a multiple of a predetermined length;

computing a hash value for the padded message;

converting the hash value into a public key encrypted message; and

embedding the public key encrypted message into the truncated image.

25. The method of claim 24 further comprising the step of maximizing spread between the verification bits.

26. The method of claim 18 further comprising the step of authenticating a symbolic representation of the graph with a text authentication algorithm.

27. The method of claim 26 further comprising the steps of:

defining nodes of the graph with specification symbols; and

defining relationships between the nodes of the graph with relationship symbols.

28. The method of claim 26 further comprising the step of coalescing the object level of the graph with the pixel level of the graph.

29. A graph authentication system comprising:

an object level authenticator for authenticating a graph at an object level;

a pixel level authenticator for authenticating the graph at a pixel level; and

an encryption system for encrypting the authenticated graph.

30. The authentication system of claim 29 wherein the object level authenticator converts the graph into a symbolic representation of the graph.

31. The authentication system of claim 30 wherein the object level authenticator includes:

a specification module for defining nodes of the graph with specification symbols;

a relationship module for defining relationships between the nodes of the graph with relationship symbols; and

a text authentication module for authenticating the symbolic representation with a text authentication algorithm.

32. The authentication system of claim 29 wherein the pixel level authenticator includes:

a visible watermarking module for adding visible authentication information to the graph; and

an invisible watermarking module for adding invisible authentication information to the graph.

33. The authentication system of claim 32 wherein the pixel level authenticator further includes a coalescing module for embedding a hash value from the object level of the graph in the pixel level of the graph.

CONTENT-BASED GRAPH AUTHENTICATION

ABSTRACT OF THE DISCLOSURE

A system and method provide for content-based authentication of binary graphs. The method includes the step of receiving an electronic file having a graphical content. An object level representation of the graphical content is then generated and authentication information is added to the electronic file based on the object level representation. The method further provides for authenticating the object level representation with a text authentication algorithm. Thus, by building a bridge from graphs to text at the character level, the present invention allows authentication of graphs using suitable text document authentication algorithms. When pixel level precision of the graph is required, a pixel level authentication can be added. This layer lets the owner detect as well as localize changes in the graph at the pixel level. Both levels of authentication are optional depending on the application.

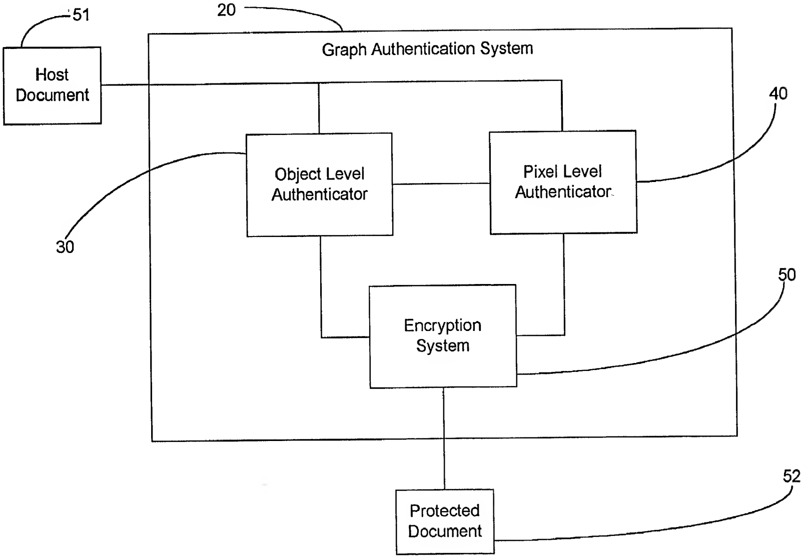

20
Graph Authentication System
51
Host Document
30
Object Level Authenticator
Pixel Level Authenticator
40
Encryption System
50
Protected Document
52

Figure 1

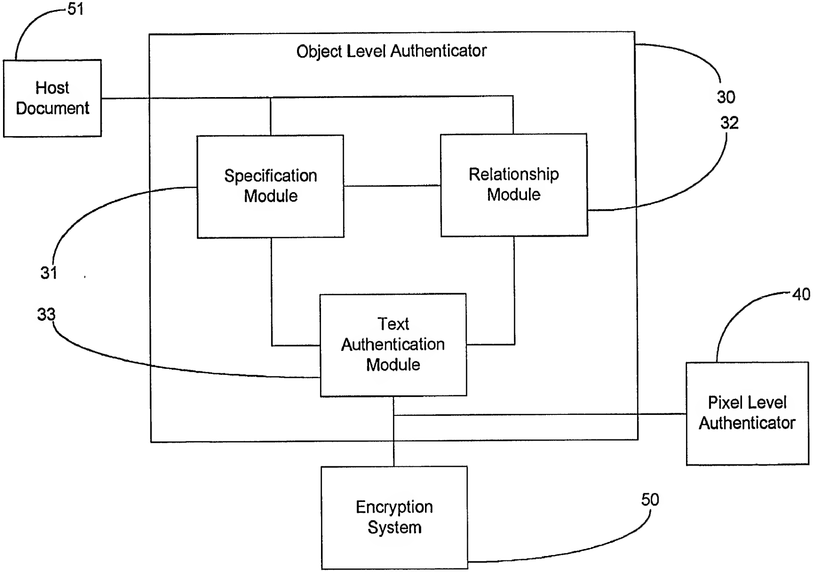

51
Host Document
Object Level Authenticator
30
32
Specification Module
Relationship Module
31
33
Text Authentication Module
40
Pixel Level Authenticator
Encryption System
50

Figure 2

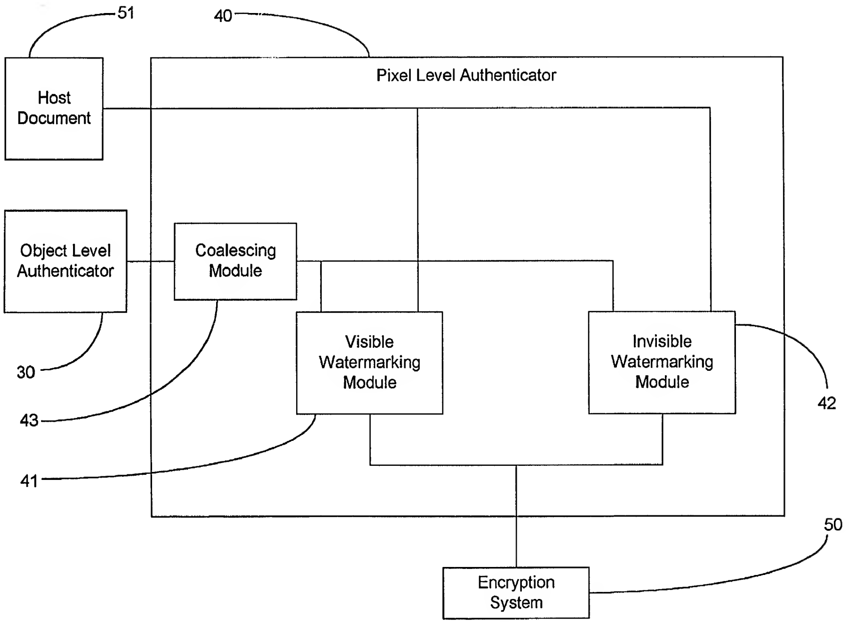

51
40
Host Document
Pixel Level Authenticator
Object Level Authenticator
Coalescing Module
Visible Watermarking Module
Invisible Watermarking Module
30
43
41
42
50
Encryption System

Figure 3

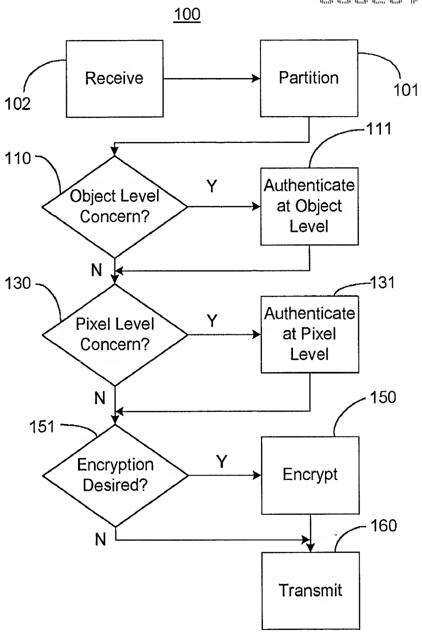


Figure 4

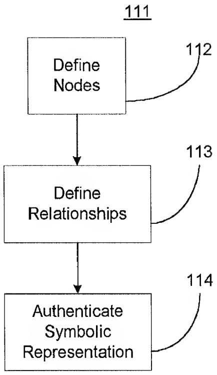


Figure 5

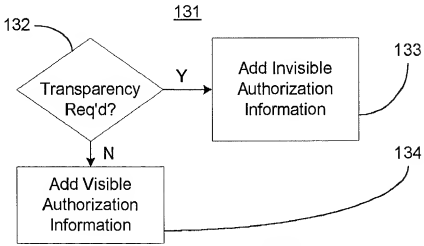

131
Transparency Req'd?
132
Y
Add Invisible Authorization Information
133
N
Add Visible Authorization Information
134

Figure 6

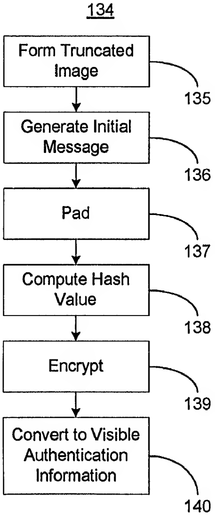


Figure 7

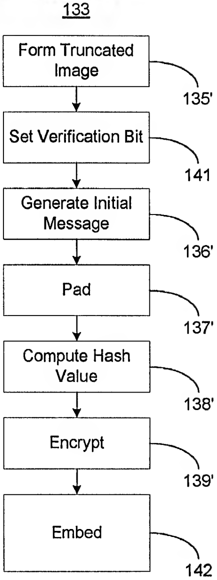


Figure 8

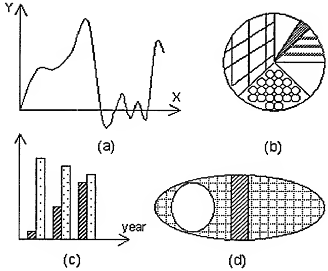


Figure 9

The system flow diagram is illustrated below. It shows the simplicity of the algorithm.

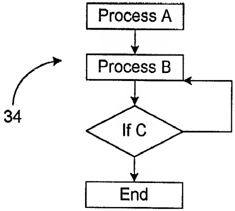


Figure 10

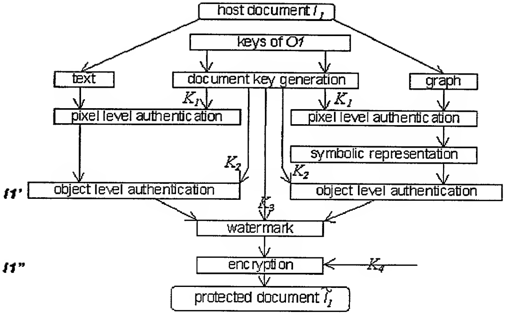

host document I_1
keys of O1
text
document key generation
graph
K_1
K_1
pixel level authentication
pixel level authentication
symbolic representation
K_2
K_2
I1'
object level authentication
object level authentication
K_3
watermark
I1"
encryption
K_4
protected document $\tilde{I}_1$

Figure 11

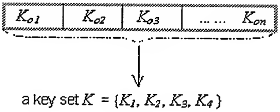


Figure 12

Relationship symbols	
< >	a tuple
∩	and
∪	or
≠	not
→	parent→child
⇒	sibling relation
⇔	twin relation
←	child←parent
>	contain relation
|	condition
.	.
.	.
.	.
:	unconnected
Specification symbols	
&	size
#	shape
@	position
©	color

Figure 13

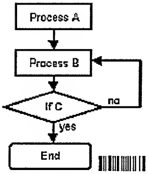
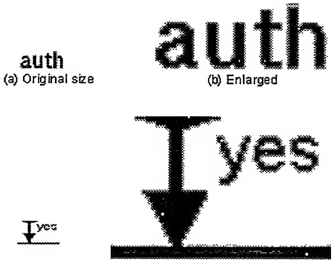


Figure 16



(a) Original size

(b) Enlarged

(c) Original size

(d) Enlarged

Figure 17

The system flow diagram is illustrated below. It shows the simplicity of the algorithm... "<N1{'Process A', #1, &reg, @mid}→N2{'Process B', #1, &reg, @mid}→N3{'If C', #3, &reg, @mid}→< N4{'End', #2, &reg, @mid}|yes; N2|no>>"

Figure 14

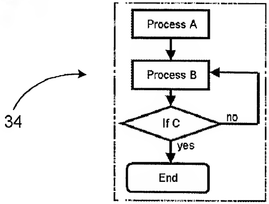


Figure 15

	W/o content-dependent one way hash		Our algorithms, w/ content-dependent one way hash		
(Text)	Traditional line spacing	Traditional serif length	Coalescing	Object level	Duel level with coalescing
Special coding	Needed	Needed	May or may not needed	May or may not needed	May or may not needed
Imperceptibility	Good	Good	OK	Good	Good
Detectability	Bad	Bad	OK	Good	Good
Pixel-level detectability	Bad	Bad	Good if Method I OK if Method II	Can't detect	OK
Localization-ability	Bad	Some bad. Some OK	OK	Good	Good
Copy and print	Bad	Bad	Good if Method I, bad if Method II	Good	Good
Noise resistance-ability	Bad	OK	Good if Method I, bad if Method II	Good	Good
Robustness to scaling	Good	OK	OK if Method I, bad if Method II	Good	Good

Figure 18

Attorney Docket No. 9432-000089

DECLARATION AND POWER OF ATTORNEY

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name,

I believe I am the original, first and sole inventor of the subject matter which is claimed and for which a patent is sought on the invention entitled

Content-Based Authentication of Graph Presented in Text Documents

the specification of which (check one)

[X] is attached hereto.

[ ] was filed on ______________________ as Application Serial No. ______________ and was amended on ______________________ (if applicable).

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to the examination of this application or to the patentability of the invention claimed therein in accordance with Title 37, Code of Federal Regulations, section 1.56.

I hereby claim foreign priority benefits under Title 35, United States Code, section 119(a)-(d) of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application on which priority is claimed:

PRIOR FOREIGN APPLICATION(S)

			Priority Claim	
(Number)	(Country)	(Day/Month/Year filed)	Yes	No
(Number)	(Country)	(Day/Month/Year filed)	Yes	No
(Number)	(Country)	(Day/Month/Year filed)	Yes	No

DECLARATION AND POWER OF ATTORNEY

I hereby claim the benefit under Title 35, United States Code, §119(e) of any United States Provisional application(s) listed below:

PRIOR PROVISIONAL APPLICATIONS

____________________ ____________________
(application serial number) (Month / Day / Year filed)

____________________ ____________________
(application serial number) (Month / Day / Year filed)

I hereby claim the benefit under Title 35, United States Code, section 120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code, section 112, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, section 1.56 which became available between the filing date of the prior application and the national or PCT international filing date of this application:

Application Serial No.	Filing Date	Status - patented, pending, abandoned

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

I hereby appoint Gregory A. Stobbs, Reg. No. 28,764, and each principal, attorney of counsel, associate and employee of Harness, Dickey & Pierce, P.L.C., who is a registered Patent Attorney, my attorney with full power of substitution and revocation, to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith. I request the Patent and Trademark Office to direct all correspondence and telephone calls relative to this application to Harness, Dickey & Pierce, P.L.C., P. O. Box 828, Bloomfield Hills, Michigan 48303 (248) 641-1600.

Full name of sole or first Inventor: Hong Heather Yu

Inventor's signature: ____________________

Date: 10/27/99

Residence: 28 Linden Ln, Plainsboro, NJ 08536, USA

Citizenship: P. R. China

Post Office Address: ____________________